Research Article

Bayesian Inference for Source Reconstruction: A Real-World Application

Eugene Yee,[1] Ian Hoffman,[2] and Kurt Ungar[2]

[1] *Defence Research and Development Canada, Suffield Research Centre, P.O. Box 4000 Stn Main,*
  Medicine Hat, AB, Canada T1A 8K6
[2] *Health Canada, Radiation Protection Bureau, 775 Brookfield Road, A.L. 6302A, Ottawa, ON, Canada K1A 1C1*

Correspondence should be addressed to Eugene Yee; eyee0309@gmail.com

This paper applies a Bayesian probabilistic inferential methodology for the reconstruction of the location and emission rate from an actual contaminant source (emission from the Chalk River Laboratories medical isotope production facility) using a small number of activity concentration measurements of a noble gas (Xenon-133) obtained from three stations that form part of the International Monitoring System radionuclide network. The sampling of the resulting posterior distribution of the source parameters is undertaken using a very efficient Markov chain Monte Carlo technique that utilizes a multiple-try differential evolution adaptive Metropolis algorithm with an archive of past states. It is shown that the principal difficulty in the reconstruction lay in the correct specification of the model errors (both scale and structure) for use in the Bayesian inferential methodology. In this context, two different measurement models for incorporation of the model error of the predicted concentrations are considered. The performance of both of these measurement models with respect to their accuracy and precision in the recovery of the source parameters is compared and contrasted.

1. Introduction

In recent years, there have been remarkable advances in sensor technology for the environmental monitoring and surveillance of chemical, biological, or radiological agents released into the atmosphere, either deliberately or accidentally. Two examples of operational monitoring sensor networks are the deployment of biological sensor arrays by the Department of Homeland Security in various cities across the United States as part of the BioWatch program [1] and the global network of radionuclide sensors that form part of the International Monitoring System deployed under the auspices of the Comprehensive Nuclear-Test-Ban Treaty [2]. In parallel with this development, there has been also significant theoretical and computational progress in the modeling and simulation of the transport and dispersion of materials released into the atmosphere, involving ever more accurate representations of the whole complex of processes on the entire spectrum of scales responsible for the dispersion of windborne materials (from the microscale turbulent motions in the atmospheric boundary layer that determine the short-range dispersion to the larger-scale motions associated with various weather systems within the whole global atmosphere that determine the long-range transport).

Lying at the nexus of these two developments is the requirement for ever more sophisticated statistical analysis tools that can be used to fuse the information embodied in the measurements of agent concentration provided by the sensors and in the predictions of these concentrations provided by the atmospheric dispersion model(s). This data-driven sensor-modeling paradigm will enable the characterization of the unknown source following event detection by the network of sensors (inverse problem), which in turn can lead to a greatly improved situational awareness and to a more informed decision-making process for response.

Inverse source modeling is a notoriously ill-posed problem in the following sense: (1) the sparse sampling (in space) of and the presence of multiple sampling times for the concentration in a dispersing cloud of contaminant may lead to difficulties of nonuniqueness in the source reconstruction

(namely, there may be many source configurations that are consistent with the limited number of concentration data); (2) the physical processes of mixing and diffusion of material in the atmosphere, with continuing dilution, smooth the concentration field leading to a progressive loss of high-frequency information that could potentially lead to an instability in the source reconstruction; and (3) actual measurements of concentration by sensors and predictions of concentrations by atmospheric dispersion model(s) are never exact (owing to measurement and model errors, resp.) which exacerbates the nonuniqueness associated with incomplete concentration observations and the instability associated with the intrinsic smoothing of the concentration field by the natural mixing processes in the atmosphere.

To overcome the problems of nonuniqueness and insta-bility in the inverse source modeling problem, the most popu-lar approach is regularization which emphasizes the selection of a "well-behaved" source configuration from the possibly infinite set of such configurations that are consistent with the limited number of noisy concentration measurements. In this approach, the idea is to find a source configuration that minimizes (optimizes) a penalty functional that involves two independent terms, namely, a regularization functional (source model norm) that imposes some constraint on the solution and a misfit functional (also some form of norm) that quantifies the discrepancy between the measured and predicted concentrations. For the application of regulariza-tion to inverse source modeling problems, see, for example, Robertson and Persson [3], Robertson and Langner [4], Thomson et al. [5], Bocquet [6], and Allen et al. [7].

Regularization provides a single "optimal" selection for the unknown source distribution (*deterministic* solution) with no rigorous quantification of the uncertainty in this par-ticular selection. An alternative to obtaining a single optimal source distribution is to characterize an ensemble of source distributions using the Bayesian statistical paradigm (pro-viding a fully *probabilistic* solution allowing the uncertainty in the source reconstruction to be rigorously assessed). The Bayesian inferential methodology for source reconstruction provides fully probabilistic information on all the parameters used to describe the unknown source distribution. The probabilistic approach using a Bayesian inferential scheme for source reconstruction has been developed by Keats et al. [8], Chow et al. [9], Senocak et al. [10], Yee et al. [11], and Yee [12–15]. The primary limiting factor in the application of the Bayesian inferential methodology for source reconstruction is computational and, as a result, it has not been as widely used as the regularization methodology.

Most of the applications of the Bayesian inferential methodology for source reconstruction have used high-quality concentration data from well-designed atmospheric dispersion experiments to validate the schema. The objective of this paper is to use concentration data obtained from an operational network of sensors (more specifically from a very small subset of the global network of radionuclide sensors that form part of the International Monitoring System) to provide a real-world test of source reconstruction based on the Bayesian statistical paradigm applied to long-range atmospheric transport on a continental or hemispheric scale.

It will be demonstrated that one of the key problems that need to be properly addressed for the successful application of Bayesian inference for source reconstruction when using operational concentration data is the "correct" specification of the expected model errors arising from the putative accuracy of a long-range forecast (or, alternatively, reanalysis) of meteorological fields and the concomitant prediction of material dispersion based on this forecast or reanalysis (which can result in significant model errors with a complex structure in the predicted concentrations required for the source inversion).

2. Bayesian Probability Theory

In a remarkable paper, Cox [16] demonstrated that proba-bility theory, when interpreted as logic, is the only calculus that conforms to a consistent theory of inference. This demonstration provides the firm logical basis for asserting that probability calculus is *the* unique quantitative theory of inference. More specifically, the cornerstone of logical inference is embodied by Bayes' theorem which itself is nothing more than the product law of probability calculus (or Bayesian probability theory):

$$p(\boldsymbol{\theta} \mid I) p(\mathbf{d} \mid \boldsymbol{\theta}, I) = p(\mathbf{d} \mid I) p(\boldsymbol{\theta} \mid \mathbf{d}, I). \qquad (1)$$

In (1), $p(\cdot)$ denotes the probability of a proposition or hypothesis, "|" denotes "conditional upon" or "given that", $\boldsymbol{\theta}$ denotes a set of parameters of a model under consideration, $\mathbf{d}$ denotes the data used for the inference, and I denotes the background assumptions or contextual information available in the problem. More specifically, as applied to the problem of source reconstruction, $\boldsymbol{\theta}$ can be identified with the set of parameters used to characterize the source distribution (e.g., location and emission rate); $\mathbf{d}$ can be associated with the measured concentration data; and I corresponds to the contextual information that is relevant to the source recon-struction problem (e.g., reanalyzed or forecast meteorology, atmospheric dispersion model used to define the source-receptor relationship).

The input quantities for Bayesian inference are on the left-hand side of (1) and are as follows: $p(\boldsymbol{\theta} \mid I)$ is the prior probability for a hypothesis about the values of the source parameter vector $\boldsymbol{\theta}$ which encodes our state of knowledge about these parameters before the receipt of the concentra-tion data $\mathbf{d}$ and $p(\mathbf{d} \mid \boldsymbol{\theta}, I)$ is the likelihood function and is considered to be a function of $\boldsymbol{\theta}$ for fixed data $\mathbf{d}$. The likelihood function incorporates the information provided by the measured concentration data $\mathbf{d}$ into the inferential scheme.

The output quantities provided by the Bayesian inference are on the right-hand side of (1) and are as follows: $p(\mathbf{d} \mid I)$ is referred to as the evidence or global likelihood and $p(\boldsymbol{\theta} \mid \mathbf{d}, I)$ is the posterior probability for the hypothesis about the values of the source parameter vector $\boldsymbol{\theta}$, evaluated in light of the additional information provided by the measured concentration data $\mathbf{d}$. The evidence is given by the following

multidimensional integral over the source parameter space [cf. (1)]:

$$p(\mathbf{d} \mid I) = \int p(\boldsymbol{\theta} \mid I)\, p(\mathbf{d} \mid \boldsymbol{\theta}, I)\, d\boldsymbol{\theta}, \tag{2}$$

which ensures the proper normalization of the posterior distribution $p(\boldsymbol{\theta} \mid \mathbf{d}, I)$ (namely, the condition that the integral of the posterior distribution over its domain of definition is unity). In the context of the determination of the plausible values for the source parameter vector (parameter estimation problem), it should be noted that the evidence is simply a normalization constant that is *independent* of the source parameter vector and, as a consequence, can therefore be ignored. However, it should be *stressed* that the evidence assumes a central role in the problem of model selection as discussed in Yee [15] and, in this context, this quantity cannot be ignored as in the parameter estimation problem considered herein.

In view of the fact that the $p(\mathbf{d} \mid I)$ is merely a normalization constant in the context of the source estimation problem, the key output of the Bayesian inference methodology in this case is the posterior probability $p(\boldsymbol{\theta} \mid \mathbf{d}, I)$ which embodies all the information about the unknown source parameters $\boldsymbol{\theta}$. In particular, $p(\boldsymbol{\theta} \mid \mathbf{d}, I)$ provides the full solution for the inverse source determination (or source term estimation) problem. Determining this quantity requires assigning appropriate functional forms for the two input quantities for Bayesian inference: (1) the prior probability $p(\mathbf{d} \mid I)$ and (2) the likelihood function $p(\mathbf{d} \mid \boldsymbol{\theta}, I)$.

2.1. Assignment of the Likelihood Function. To assign a specific functional form for the likelihood function, one needs to relate the source parameters $\boldsymbol{\theta}$ to the available concentration data $\mathbf{d}$. In this sense, the likelihood function defines the probabilistic model of how the concentration data were generated. Towards this purpose, if the concentration data acquired (by a sensor) at the space-time point $(\mathbf{x}_J, t_J)$ is represented by $d(\mathbf{x}_J, t_J)$, then the concentration data and the source parameters $\boldsymbol{\theta}$ are assumed to be related by

$$\begin{aligned} d_J &\equiv d(\mathbf{x}_J, t_J) \\ &= \overline{C}(\boldsymbol{\theta}; \mathbf{x}_J, t_J) + e(\mathbf{x}_J, t_J) \\ &\equiv \overline{C}_J(\boldsymbol{\theta}) + e_J, \quad J = 1, 2, \ldots, N, \end{aligned} \tag{3}$$

where N is the total number of measured concentration data. In (3), $\overline{C}_J(\boldsymbol{\theta})$ is the modeled (predicted) mean concentration at the Jth space-time point. The predicted concentration is determined using an atmospheric dispersion model for a source distribution characterized by the parameter vector $\boldsymbol{\theta}$.

The error (discrepancy) between the measured concentration d_J and the predicted concentration $\overline{C}_J(\boldsymbol{\theta})$ is represented symbolically in (3) by e_J. In the problem considered in this paper, there are two major contributions to the error e_J, namely, the observation or instrument error in the measured concentration d_J arising from the noise inherent in the sensor and the model error in the determination of the predicted concentration $\overline{C}_J(\boldsymbol{\theta})$. Of these two errors, the model error is

by far the most dominant contribution to the total error e_J and is also the most difficult to characterize.

In our current application, the model error arises from three primary sources. These are as follows:

(1) uncertainties in the representation of various physical processes in the dispersion model;

(2) uncertainties in the input meteorological fields (initial and boundary conditions) used to "drive" the dispersion model (either numerical weather prediction uncertainties if these fields are obtained as a forecast or data assimilation uncertainties for the state of the atmosphere if these fields are obtained through a reanalysis);

(3) uncertainties in the numerical solution of the model equations that characterize the dispersion model which comprise both discretization (including representativity) errors and statistical model errors, the latter of which arises from using necessarily a finite number of "marked" fluid particles to estimate the mean concentration field in the case of a Lagrangian stochastic model of dispersion.

As a consequence of the complexity in structure of the error e_J [cf. (3)] for our current application, it is extremely difficult (if not insuperable) to specify *a priori* an exact value σ_J for the standard deviation of e_J ($J = 1, 2, \ldots, N$). If the standard deviation σ_J of e_J was exactly known, then it can be shown by application of the principle of maximum entropy that a Gaussian distribution of the form

$$p(\mathbf{d} \mid \boldsymbol{\theta}, I) = \frac{1}{\prod_{J=1}^{N} \sqrt{2\pi}\sigma_J} \exp\left(-\frac{1}{2}\chi^2(\boldsymbol{\theta})\right), \tag{4}$$

where

$$\chi^2(\boldsymbol{\theta}) \equiv \sum_{J=1}^{N} \left(\frac{d_J - \overline{C}_J(\boldsymbol{\theta})}{\sigma_J}\right)^2, \tag{5}$$

would be the most conservative choice for the direct probability (or likelihood) of the concentration data $\mathbf{d} \equiv (d_1, d_2, \ldots, d_N)$ [17].

Because σ_J is not known *a priori*, it is useful to characterize the uncertainty in the specification of σ_J with a probability distribution. Following Yee [15], we choose this probability distribution to be an inverse gamma distribution with the following form:

$$\varphi(\sigma_J \mid s_J, \alpha, \beta) = 2\frac{\alpha^\beta}{\Gamma(\beta)}\left(\frac{s_J}{\sigma_J}\right)^{2\beta} \exp\left(-\alpha\frac{s_J^2}{\sigma_J^2}\right)\frac{1}{\sigma_J}, \tag{6}$$

$$J = 1, 2, \ldots, N.$$

Here, $\Gamma(x)$ denotes the gamma function, α and β are scale and shape parameters of the inverse gamma distribution, and s_J is the quoted (nominal) estimate for the true but unknown standard deviation σ_J. The values for the hyperparameters α and β are chosen as $\alpha = \pi^{-1}$ and $\beta = 1$ following the rationale described in Yee [15].

The likelihood function in (4) and (5) depends on the error standard deviations σ_J ($J = 1, 2, \ldots, N$) which are generally unknown. To remove these unwanted parameters (nuisance parameters), we can multiply the likelihood given in (4) by the (assigned) probability distribution for each of the error standard deviations embodied in (6) and integrate the result with respect to the unwanted parameters (error standard deviations) to give an integrated likelihood function with the following form [15]:

$$
\begin{aligned}
& p\left(\mathbf{d} \mid \boldsymbol{\theta}, \mathbf{s}, \alpha, \beta\right) \\
& = \int p\left(\mathbf{d} \mid \boldsymbol{\theta}, I\right) \prod_{J=1}^{N} \varphi\left(\sigma_J \mid s_J, \alpha, \beta\right) d\boldsymbol{\sigma} \\
& = \prod_{J=1}^{N} \frac{\alpha^{\beta} \Gamma\left(\beta + 1/2\right)}{\sqrt{2\pi} s_J \Gamma\left(\beta\right)} \left[\alpha + \frac{\left(d_J - \overline{C}_J\left(\boldsymbol{\theta}\right)\right)^2}{2 s_J^2}\right]^{-\beta - 1/2}.
\end{aligned}
\tag{7}
$$

In (7), $\mathbf{s} \equiv (s_1, s_2, \ldots, s_N)$ is the vector of estimated (quoted) standard deviations for the error e_J ($J = 1, 2, \ldots, N$) and $d\boldsymbol{\sigma} \equiv d\sigma_1 d\sigma_2 \ldots d\sigma_N$.

To be more specific with respect to the source parameter vector $\boldsymbol{\theta}$, a source distribution S with the following form is considered in this paper:

$$
S(\mathbf{x}) = Q_s \delta(\mathbf{x} - \mathbf{x}_s),
\tag{8}
$$

where $\delta(\cdot)$ is the Dirac delta function. Equation (8) describes a (continuous) point source located at the vector position $\mathbf{x}_s$ that is emitting contaminant at a constant emission rate Q_s. The parameters describing this source distribution can be assembled into a source parameter vector given by $\boldsymbol{\theta} \equiv (\mathbf{x}_s, Q_s)$.

2.2. Assignment of the Prior. To assign the prior probability for the source parameters $\boldsymbol{\theta}$, it is necessary to state explicitly what is known about these parameters. Firstly, it is assumed that the source parameters are logically independent with the result that the prior distribution $p(\boldsymbol{\theta} \mid I)$ factorizes as follows:

$$
p(\boldsymbol{\theta} \mid I) = p(\mathbf{x}_s \mid I) p(Q_s \mid I).
\tag{9}
$$

Secondly, the source location and emission rate are known *a priori* to be bounded. In particular, it is assumed that the location $\mathbf{x}_s$ of the source is contained in some spatial region $\mathcal{D} \subset \mathbb{R}^3$. Furthermore, the emission rate Q_s is assumed to be bounded by $Q_{\min} < Q_s < Q_{\max}$ where $Q_{\min}$ and $Q_{\max}$ are the lower and upper bounds for the emission rate, respectively. Finally, if nothing else is known about these parameters except for the bounds, then application of the principle of maximum entropy to our state of knowledge concerning the source parameters results in the assignment of a uniform prior distribution for these parameters, so

$$
p(\boldsymbol{\theta} \mid I) \propto \mathbb{I}_{\mathcal{D}}(\mathbf{x}_s) \mathbb{I}_{(Q_{\min}, Q_{\max})}(Q_s),
\tag{10}
$$

where $\mathbb{I}_A(x)$ denotes the indicator function for set A defined as $\mathbb{I}_A(x) = 1$ if $x \in A$ and $\mathbb{I}_A(x) = 0$ if $x \notin A$.

2.3. Posterior Distribution: The Output. The key output of the Bayesian inference is the posterior distribution $p(\boldsymbol{\theta} \mid \mathbf{d}, I)$ which can be obtained by combining the assignments for the likelihood function and the prior distribution given by (7) and (10), respectively, to give [cf. (1)]

$$
\begin{aligned}
& p\left(\boldsymbol{\theta} \mid \mathbf{d}, \mathbf{s}, \alpha, \beta, I\right) \\
& \quad \propto \mathbb{I}_{\mathcal{D}}\left(\mathbf{x}_s\right) \mathbb{I}_{(Q_{\min}, Q_{\max})}\left(Q_s\right) \\
& \quad \times \prod_{J=1}^{N} \frac{\alpha^{\beta} \Gamma\left(\beta + 1/2\right)}{\sqrt{2\pi} s_J \Gamma\left(\beta\right)} \\
& \quad \times \left[\alpha + \left(d_J - \overline{C}_J(\boldsymbol{\theta})\right)^2 / (2 s_J^2)\right]^{-\beta - 1/2}.
\end{aligned}
\tag{11}
$$

Note that the quantities $\mathbf{s}$, α, and β have been added explicitly to the posterior probability of the source parameters $\boldsymbol{\theta}$ in (11) to indicate that these quantities are known (namely, they are provided *a priori* by the user, in addition to the measured concentration data $\mathbf{d}$).

3. Computational Aspects

There are two computational problems that need to be addressed before the Bayesian inferential methodology for source reconstruction can be applied to practical real-world applications; namely, (1) that Bayesian inversion of concentration data requires a fast and efficient technique for the determination of the source-receptor relationship (namely, for the fast computation of $\overline{C}_J(\boldsymbol{\theta})$ for a given hypothesis about the source parameters $\boldsymbol{\theta}$) and (2) methodology for the efficient sampling of the posterior distribution $p(\boldsymbol{\theta} \mid \mathbf{d}, \mathbf{s}, \alpha, \beta, I)$.

To address the first problem, Keats et al. [8] and Yee et al. [11] demonstrated that the use of a receptor-oriented representation for the source-receptor relationship (rather than the more usual source-oriented representation) enabled the efficient computation of the predicted concentration at a fixed receptor location associated with the exploration of a large number of source parameter hypotheses required in the simulation-based Bayesian inference methodology. For the application reported herein, a backward Lagrangian stochastic (LS) model for long-range transport was used to determine the adjoint concentration field C^* over the northern hemisphere.

The backward LS model employed here is an operational model used by the Canadian Meteorological Centre to support both Canadian Treaty monitoring and the various mandates of the Provisional Technical Secretariat of the Comprehensive Nuclear-Test-Ban Treaty Organization, including event analysis by member states. The backward LS model for the determination of C^* was "driven" by reanalyzed meteorological fields that were obtained at a relatively low temporal and spatial resolution, namely, at a temporal resolution of 6 h (rate of the data assimilation) with a core spatial resolution of 0.5° on a geographical latitude and longitude coordinate system. The backward LS model was run retrospectively using these reanalyzed meteorological fields, providing C^* fields with a temporal resolution of 3 hours over

a period of 14 days prior to the commencement of the sampling for a particular activity concentration measurement.

It should be noted that, with the source distribution S given by (8), the predicted concentration at a receptor location $\mathbf{x}_J$ and time t_J can be easily determined from the following relationship once the adjoint concentration field C^* has been computed for this location and time:

$$
\begin{aligned}
\overline{C}_J(\boldsymbol{\theta}) &\equiv \overline{C}(\boldsymbol{\theta}; \mathbf{x}_J, t_J) \\
&= \int_{-\infty}^{t_J} dt' \int_{\mathscr{D}} d\mathbf{x}' C^*\left(\mathbf{x}', t' \mid \mathbf{x}_J, t_J\right) S\left(\mathbf{x}'\right) \\
&\equiv \langle C^* \mid S \rangle (\mathbf{x}_J, t_J),
\end{aligned}
\tag{12}
$$

where $C^*(\mathbf{x}', t' \mid \mathbf{x}_J, t_J)$ is the adjoint concentration at space-time point $(\mathbf{x}', t')$ associated with the sensor concentration datum measured at location $\mathbf{x}_J$ and time t_J. In (12), $\mathscr{D}$ is a volume in space that contains the source. The reader is referred to Marchuk [18] and Le Dimet and Talagrand [19] for a more detailed explanation of the adjoint formulation in a more general context.

Note that the predicted mean concentration $\overline{C}_J(\boldsymbol{\theta})$ "seen" by a sensor for a given hypothesis about the source parameters $\boldsymbol{\theta}$ can be rapidly computed by simply evaluating the inner (or scalar) product $\langle C^* \mid S \rangle$ of the adjoint concentration C^* and the source distribution S corresponding to the given hypothesis. In other words, the predicted concentration $\overline{C}_J$ is obtained from a mathematical model (backward LS model for dispersion) by evaluation of the bounded linear functionals $\overline{C}_J = \langle C_J^* \mid S \rangle$ for $J = 1, 2, \ldots, N$ where C_J^* denotes the adjoint concentration field obtained at the sensor space-time point $(\mathbf{x}_J, t_J)$ and where S corresponds to a given hypothesis about the source. In applied mathematics, the elements C_J^* are referred to usually as representers. More specifically, if we substitute (8) into (12), the model (predicted) concentration $\overline{C}_J(\boldsymbol{\theta})$ "seen" by the sensor at location $\mathbf{x}_J$ and time t_J is given explicitly by

$$
\overline{C}_J(\boldsymbol{\theta}) = Q_s \int_{-\infty}^{t_J} C^*\left(\mathbf{x}_s, t' \mid \mathbf{x}_J, t_J\right) dt'.
\tag{13}
$$

To address the second computational problem mentioned above, we apply a Markov chain Monte Carlo (MCMC) algorithm for the efficient sampling of the posterior distribution of the source parameters (see Gilks et al. [20], Gelman et al. [21], Savchuk and Tsokos [22], and Yuen [23]). In general, MCMC algorithms generate a Markov chain whose stationary distribution coincides exactly with the target probability distribution that we are trying to sample from (which in our case is the posterior distribution $p(\boldsymbol{\theta} \mid \mathbf{d}, \mathbf{s}, \alpha, \beta, I)$ given in (11)). The Metropolis-Hastings (M-H) algorithm [20, 21] forms the underlying basis for MCMC sampling and it is perhaps not too surprising that the M-H algorithm has become almost synonymous with MCMC sampling. Indeed, most of the algorithms for MCMC sampling reported in the literature [20–23] can be interpreted as either special cases or extensions of the basic M-H algorithm.

For the current study, we use a multiple-try differential evolution adaptive Metropolis algorithm with sampling from an archive of past states which is referred to as MT-DREAM$_{(ZS)}$. The details of this MCMC sampling algorithm are described by Laloy and Vrugt [24], but, for completeness, we will briefly summarize the main components of this algorithm. In particular, only the relevant details of the algorithm that are required for the interpretation of the results in this paper are emphasized. Firstly, MT-DREAM$_{(ZS)}$ samples from an archive of past states to generate the candidate points (proposals) for each of the individual Markov chains that are used to explore the target posterior distribution. Secondly, as already alluded to here, the algorithm utilizes multiple (different) Markov chains that are run simultaneously in parallel. These multiple chains employ a self-adaptive randomized subspace sampling of difference vectors from the archive of past states to generate new candidate points in these chains. As part of the randomized subspace sampling strategy, each element of the candidate points for the parallel proposals is updated (accepted) in accordance with a binomial scheme (Bernoulli trial) with a crossover probability p_s; otherwise, the proposed element retains its previous (old) value. This multiple-chain approach automatically adjusts or adapts the scale and orientation of the proposal function. Thirdly, a snooker step update of the state with an adaptive step size is included with a fixed (albeit small) probability in order to improve the mixing efficiency of the algorithm for exploration of the hypothesis space. The updated states from the various parallel Markov chains are periodically stored in the archive of past states after every K ($K \geq 1$) updates.

Fourthly, to further improve the efficiency of the sampling, MT-DREAM$_{(ZS)}$ incorporates a multiple-try Metropolis (MTM) approach proposed initially by Liu et al. [25]. The basic idea underpinning the MTM approach is as follows: longer range candidate moves are rarely accepted, but if multiple points are proposed for these longer range moves then the acceptance probability will be increased. The MTM algorithm is applied individually to each of the different Markov chains used in the MT-DREAM$_{(ZS)}$, involving generating K draws using the randomized subspace sampling procedure for each chain, choosing one of these draws (proposals) as the reference point, and generating a new set of $(K - 1)$ draws with respect to this reference point using the randomized subspace sampling strategy. The acceptance rule is simply the Metropolis-Hasting acceptance probability applied to the sequence of proposals that comprise the MTM schema.

Finally, to determine if the multiple Markov chains used in MT-DREAM$_{(ZS)}$ have achieved stationarity (namely, have converged to the stationary distribution associated with the chains), the Gelman and Rubin [26] convergence diagnostic $\widehat{R}$ is computed for each dimension of each chain using the last 50% of the samples of the chain. This simple convergence statistic determines whether a chain has achieved stationarity by comparing the variance within each chain and the variance between chains (interchain variance).

4. Example: A Real-World Application

The International Monitoring System (IMS) consists of a comprehensive network of seismic, hydroacoustic,

FIGURE 1: Locations of the three sampling stations from the International Monitoring System radionuclide network used for the source reconstruction. The release location of the Xe-133 tracer source was at Chalk River Laboratories.

infrasound, and radionuclide sensors as part of the verification regime of the Comprehensive Nuclear-Test-Ban Treaty (CTBT) which bans nuclear explosions. A subset of the IMS is the subnetwork of radionuclide gamma detectors/particle filters for the measurement of the activity concentration for various radionuclides (e.g., particulate or aerosol species such as Cs-137 and I-131 and/or noble gases such as Xe-133). The IMS radionuclide network will (eventually) have 80 monitoring stations worldwide for the measurement of the activity concentration for particulate/aerosol radioactive species, of which at least 40 stations would also have the capability to measure the activity concentration of noble gases [2]. The stations provide 12 or 24 h averaged activity concentrations of the various radionuclides depending on the technology used.

We applied the proposed source reconstruction methodology to some (albeit limited) measurements of activity concentrations by a very small subset of the IMS radionuclide network. The release event consisted of the stack emissions from Chalk River Laboratories (CRL) which houses an international production facility for medical radioisotopes. Chalk River Laboratories, which at a latitude of 46.15°N and at a longitude of -77.37°E, is located about 180 km northwest of the city of Ottawa, Ontario. A characterization of the weekly stack emissions of Xe-133 from the CRL medical isotope production facility over a 5-year period yielded a median daily emission of about 24 TBq (or, equivalently, an emission rate of about 1.0×10^{12} Bq h^{-1}).

Xe-133 activity concentrations were detected and measured by three sampling sites in North America. These three sampling sites are shown in Figure 1. The three sites displayed are part of the noble gas monitoring network of the CTBT verification regime. More specifically, the three stations are as follows: CAX17 (St. John's, Newfoundland, at latitude 47.59°N and longitude -52.74°E); USX75 (Charlottesville, Virginia, at latitude 38.0°N and longitude -78.4°E); and USX74 (Ashland, Kansas, at latitude 31.17°N and longitude -99.77°E).

At each monitoring site, one of two different monitoring technologies is employed to measure radioxenon. The St. John's site has a Système de Prélèvements et d'Analyse en Ligne d'Air pour Quantifier le Xénon (SPALAX) high-resolution gamma system operating on a 24 h sample collection period, while the remaining sites have a Swedish Automatic Unit for Noble Gas Acquisition (SAUNA) beta-gamma coincidence system with a 12 h sample collection period. Both systems employ activated charcoal to remove xenon from the air for radioxenon analysis. After the measurement process is complete, the xenon sample can be stored in an archive bottle for optional remeasurement either on-site or in an off-site laboratory. The stable xenon volume collected varies approximately from 2 mL to 8 mL depending on the technology used, but both technologies have a roughly equivalent Xe-133 sensitivity of approximately 0.2 mBq m^{-3}.

The monitoring data for activity concentrations of Xe-133 were obtained from a single month (December 2011). The three monitoring stations used for this study were known to be operating normally during this period of time. For the purposes of source reconstruction, we used 36 of these concentration samples extracted from the three sampling sites. The concentration time series data at each station were blocked averaged over a temporal duration of 3 days to give 8 concentration data d_J that were used as the input for the source reconstruction methodology (namely, $N = 8$ for this example). It should be noted that predictions for the temporally averaged concentration data were subject to smaller statistical sampling errors (for a fixed number of marked fluid particles used in the backward LS model), which is the primary reason for aggregating the 12 or 24 h measurements in this manner. As mentioned earlier, reanalyzed meteorological fields were used to "drive" an operational backward LS model to determine C^* which were utilized subsequently for the fast calculation of the predicted concentration $\overline{C}_J(\boldsymbol{\theta})$ for an arbitrary hypothesis $\boldsymbol{\theta}$ about the source.

In addition to the measured activity concentration data d_J and the predicted activity concentration data $\overline{C}_J(\boldsymbol{\theta})$, it is also necessary to provide an estimate for the noise error variance s_J^2 [cf. (7)]. The estimate for the noise error variance includes two major contributions; namely, (1) an estimate for the sensor error sampling variance $s_{e,J}^2$ in the measurement of d_J and (2) an estimate for the model error variance $s_{m,J}^2$ in the prediction of $\overline{C}_J(\boldsymbol{\theta})$. These estimates for the two contributions to the noise error variance were added in quadrature to give $s_J^2 = s_{e,J}^2 + s_{m,J}^2$. Very good estimates for the sensor error standard deviation (or square root of the variance) were provided for the expected precision in the measurements of the activity concentration at each of the three sampling stations. In contrast, no estimates for the precision in the predicted concentrations were available. In light of this, the model error standard deviation was assumed (rather arbitrarily) to be 50% of the predicted concentration $\overline{C}_J(\boldsymbol{\theta})$.

For our example, the emission source is assumed to be at or near ground level so that the height of the source is not of any interest in the reconstruction. As a result, the unknown
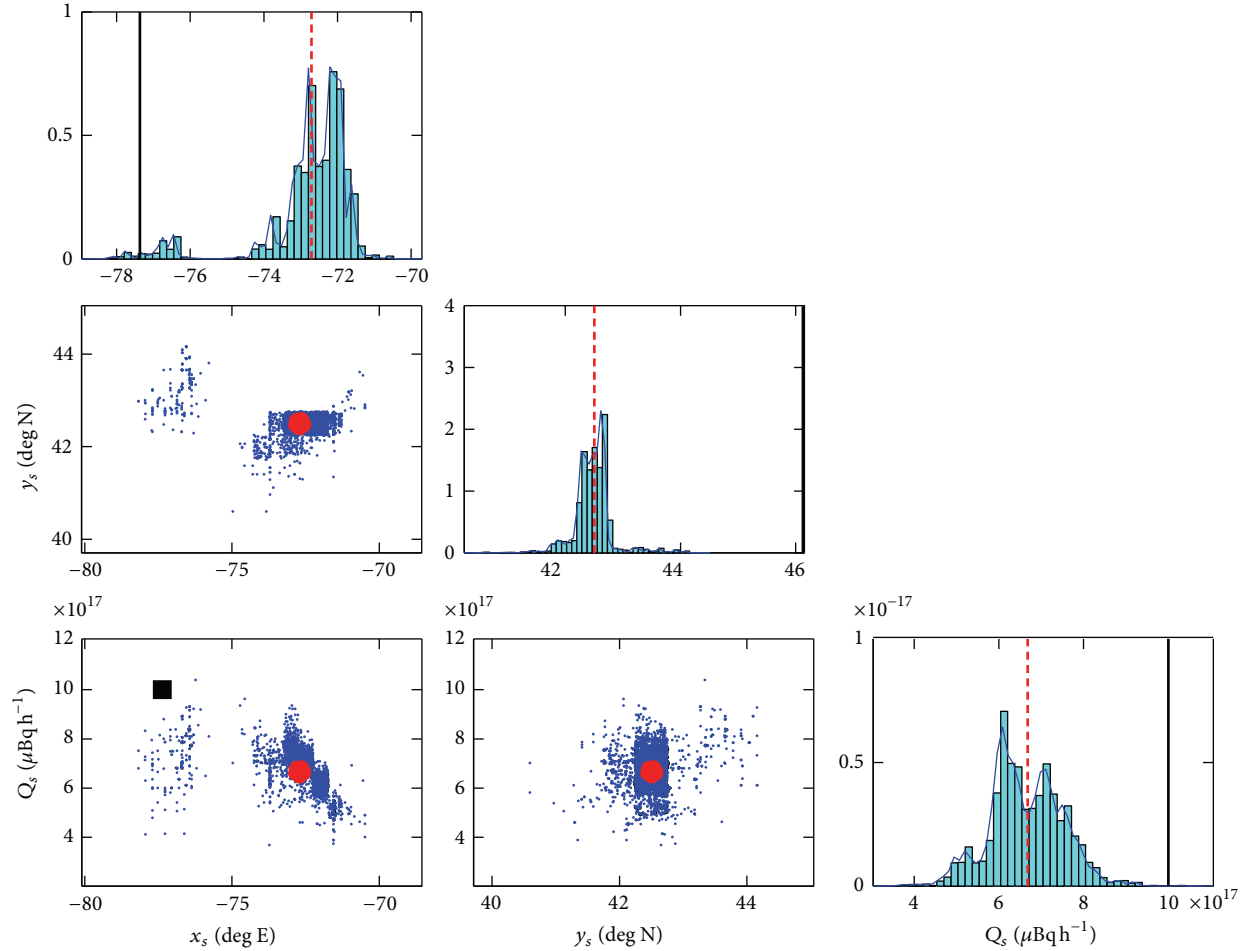
FIGURE 2: Univariate (diagonal) and bivariate (off-diagonal) marginal posterior distributions of the source parameters; namely, longitudinal position x_s, latitudinal position y_s, and emission rate Q_s. The true parameter values are shown by a solid square or a solid vertical line and the best estimates of the parameter values are represented as a solid circle or a dashed vertical line.

location parameters for the source are its longitudinal (x_s) and latitudinal (y_s) positions, so $\boldsymbol{\theta} = (x_s, y_s, Q_s)$. The proposed stochastic sampling algorithm constructs an initial population for the archive of past states by sampling from the prior distribution $p(\boldsymbol{\theta} \mid I)$ (see (10)). The hyperparameters defining the prior distribution $p(\boldsymbol{\theta} \mid I)$ are chosen as follows: $Q_{min} = 1.0 \times 10^{15}\,\mu\text{Bq h}^{-1}$, $Q_{max} = 1.0 \times 10^{20}\,\mu\text{Bq h}^{-1}$ which prescribes the lower and upper bounds for the source emission rate and $\mathcal{D} = (-125°\text{E}, -45°\text{E}) \times (25°\text{N}, 75°\text{N})$ which defines the prior bounds for the location (x_s, y_s) of the source. It is noted that $\mathcal{D}$ encloses the North American continent.

The samples of $\boldsymbol{\theta}$ drawn from the posterior distribution $p(\boldsymbol{\theta} \mid \mathbf{d}, \mathbf{s}, \alpha, \beta, I)$ (cf. (11)) using the MT-DREAM$_{(ZS)}$ sampling algorithm were used to determine the characteristics (location and emission rate) of the source. Estimates of these source parameters were obtained directly from the multiple chains of samples generated by the MCMC algorithm after convergence of the chains as determined by the Gelman-Rubin $\widehat{R}$ statistic. Figure 2 exhibits the univariate (diagonal) and bivariate (off-diagonal) marginal posterior distributions for the source parameters. For the univariate marginal

posterior distribution of a parameter, the solid vertical line delineates the true value of the parameter, and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean. Similarly, for the bivariate marginal posterior distribution of various combinations of parameters, the solid square marks the position of the true source parameter values and the solid circle exhibits the best estimate of the true source parameter values obtained as the posterior mean. It should be noted that the axes limits are chosen to display the regions of the highest probability in the marginal posterior probability distributions for the parameters and, in certain cases, these regions do not contain the true parameter values (with the result that some of the panels in Figure 2 do not contain either the solid square or solid vertical line representing the true parameter values).

The posterior mean, posterior standard deviation, and lower and upper bounds for the 95% highest posterior distribution (HPD) interval (the $p\%$ HPD (or credible) interval is that interval that contains the parameter with a $p\%$ (posterior) probability, with the lower and upper bounds of the interval specified such that the probability density within the interval is everywhere larger than that outside it) for the

TABLE 1: The posterior mean, posterior standard deviation, and lower and upper bounds of the 95% HPD interval of the parameters x_s (°E), y_s (°N), and Q_s (μBq h^{-1}) calculated from samples of $\boldsymbol{\theta}$ drawn from the posterior distribution $p(\boldsymbol{\theta}|\mathbf{d}, \mathbf{s}, \alpha, \beta, I)$.

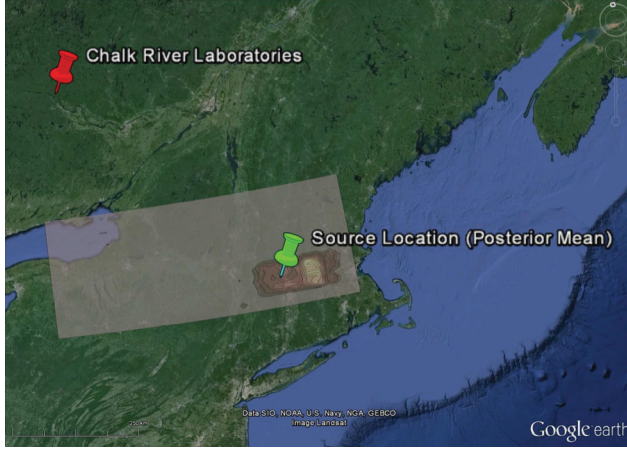| Parameter	Mean	Standard deviation	95% HPD	Actual
x_s (°E)	−72.71	1.22	(−76.71, −71.56)	−77.37
y_s (°N)	42.51	0.32	(41.86, 43.35)	46.15
Q_s (μBq h^{-1})	6.68×10^{17}	8.55×10^{16}	$(4.98, 8.33) \times 10^{17}$	1.0×10^{18}

FIGURE 3: Two-dimensional marginal posterior distribution of the source location georeferenced on a Google Earth image.

reconstructed source parameters are summarized in Table 1. It is noted that the accuracy in the reconstruction of the source parameters is fairly good for both the location (considering the fact that the reconstruction was undertaken on a continental scale) and the emission rate. More specifically, the distance between the true source location and the best estimate of the source location (obtained as the posterior mean) is only about 572 km (see Figure 3) and the best estimate of the emission rate (obtained again as the posterior mean) is within 33% of the true emission rate.

Even so, a perusal of Table 1 indicates that the precision in the estimates of both the source location and the emission rate is poorly determined. Indeed, it should be noted that the 95% HPD intervals for longitudinal/latitudinal positions and the emission rate of the reconstructed source do not contain the true source parameters. This defect in the source reconstruction can be attributed to the difficulties in providing good estimates for the model errors in the prediction of $\overline{C}_J(\boldsymbol{\theta})$. It is evident that the inability to provide good estimates for the model errors (which in principle can be quite significant in our example) can lead to a loss of power in the source reconstruction and may mask important features of the measured activity concentration data.

Given the complexity in the heteroskedastic variance of the model error in our application, the specification of the prior information on the structure of the model error is extremely difficult. In light of these seemingly insuperable difficulties, we consider an alternative measurement model

to that introduced earlier in (3). To this purpose, let us now focus on the following alternative measurement model:

$$d_J = m_J \overline{C}_J(\boldsymbol{\theta}) + n_J, \quad J = 1, 2, \ldots, N, \tag{14}$$

where m_J are unknown multipliers (scale factors) that are applied to the predicted concentration $\overline{C}_J(\boldsymbol{\theta})$ in order to compensate for the model uncertainty. It is important to note that, in (14), the n_J term represents *only* the measurement error in the activity concentration d_J. The model errors incurred in the predicted concentration $\overline{C}_J(\boldsymbol{\theta})$ are compensated by the introduction of the multiplier m_J.

For the alternative measurement model, the multipliers m_J ($J = 1, 2, \ldots, N$) are unknown parameters and need to be estimated in addition to the source parameters. Let us denote the source parameters in this alternative model by $\boldsymbol{\theta}^s \equiv (\mathbf{x}_s, Q_s)$ and by $\boldsymbol{\theta}^m \equiv (m_1, m_2, \ldots, m_N)$ all other relevant parameters (multipliers in our example). These latter parameters are usually referred to as nuisance parameters. Both sets of parameters define the parameter vector as $\boldsymbol{\theta} = (\boldsymbol{\theta}^s, \boldsymbol{\theta}^m)$. The likelihood function for the alternative measurement model is still given by (7) with the implicit understanding that the estimated noise variances s_J^2 now *only* include the measurement error contribution (namely, $s_J^2 = s_{e,J}^2$). As already noted above, this contribution to the uncertainty is well characterized in our current application (implying that the prior uncertainty in the measurement errors can be specified correctly). For the alternative measurement model, one requires also the specification of the prior distributions for the multipliers (nuisance parameters). Towards this objective, uniform priors defined over the range from $m_{\min}$ to $m_{\max}$ will be adopted as priors for the multipliers. In view of this, the prior distribution for the alternative measurement model replaces (10) by the following assignment:

$$p(\boldsymbol{\theta} \mid I) \propto \mathbb{I}_{\mathcal{D}}(\mathbf{x}_s) \, \mathbb{I}_{(Q_{\min}, Q_{\max})}(Q_s) \times \prod_{J=1}^{N} \mathbb{I}_{(m_{\min}, m_{\max})}(m_J). \tag{15}$$

Using the new measurement model and the modified likelihood function and prior distribution, we applied the MT-DREAM$_{(ZS)}$ algorithm to sample from the modified posterior distribution for $\boldsymbol{\theta}$. The hyperparameters that define the prior distribution for the source parameters $\boldsymbol{\theta}^s$ were exactly as described above. The prior bounds for the multipliers m_J were $m_{\min} = 0.1$ and $m_{\max} = 10.0$ for $J = 1, 2, \ldots, 8$ (recall that $N = 8$ in our example). The one-dimensional and two-dimensional marginal posterior distributions of the

TABLE 2: The posterior mean, posterior standard deviation, and lower and upper bounds of the 95% HPD interval of the parameters x_s (°E), y_s (°N), and Q_s (μBq h^{-1}) calculated from samples of θ drawn from the posterior distribution $p(\theta|\mathbf{d}, \mathbf{s}, \alpha, \beta, I)$. The results are obtained using the alternative measurement model which employs multipliers to compensate for the unknown model error structure.

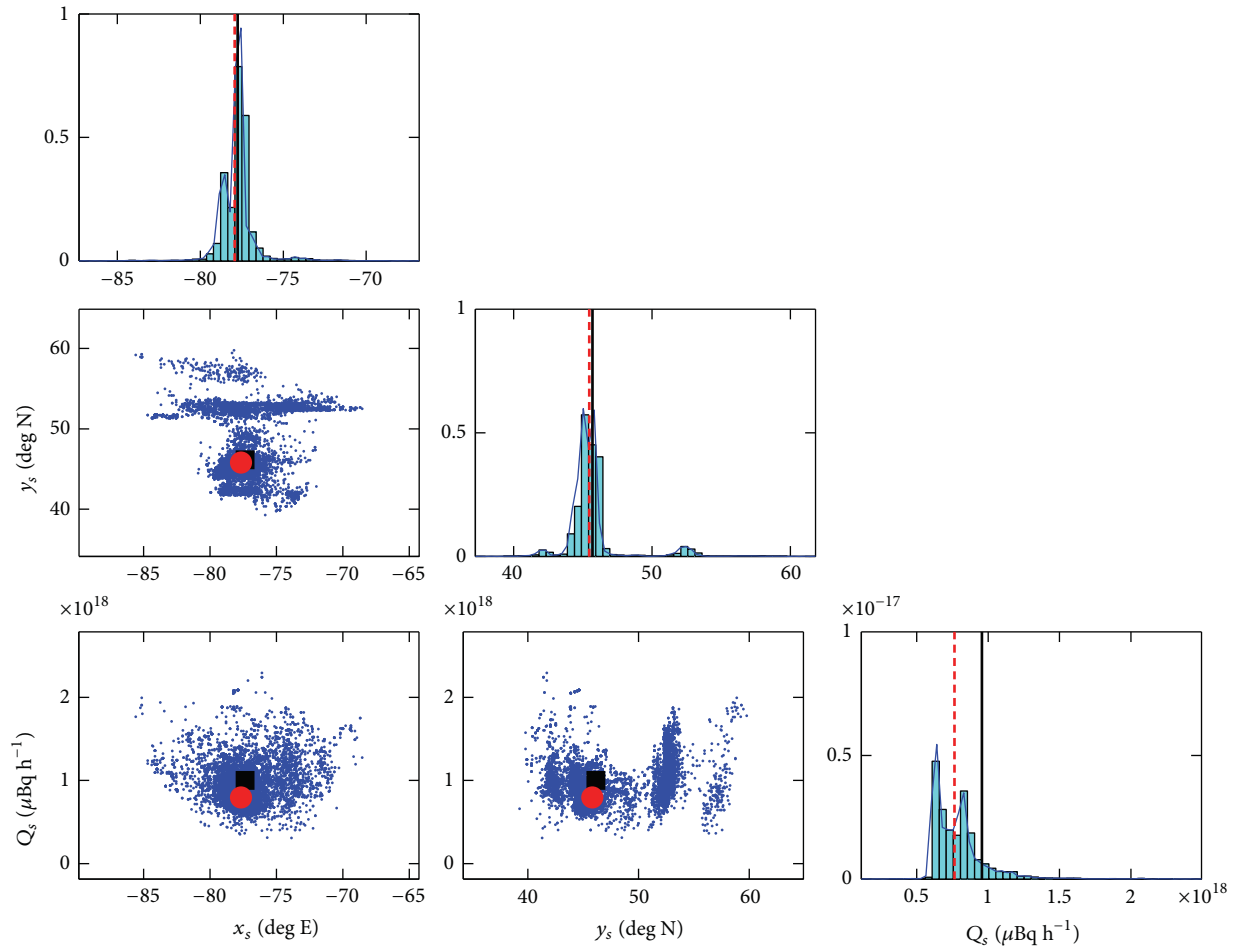| Parameter	Mean	Standard deviation	95% HPD	Actual
x_s (°E)	−77.66	1.04	(−79.21, −74.88)	−77.37
y_s (°N)	45.80	1.96	(42.69, 52.64)	46.15
Q_s (μBq h^{-1})	7.97×10^{17}	1.75×10^{17}	$(6.37, 12.5) \times 10^{17}$	1.0×10^{18}

FIGURE 4: Univariate (diagonal) and bivariate (off-diagonal) marginal posterior distributions of the source parameters; namely, longitudinal position x_s, latitudinal position y_s, and emission rate Q_s. The true parameter values are shown by a solid square or a solid vertical line and the best estimates of the parameter values are represented as a solid circle or a dashed vertical line. The reconstruction was obtained using the alternative measurement model which employs multipliers to compensate for the unknown model error structure.

source parameters θ^s are displayed in Figure 4. The true source parameters represented by either the solid square or solid vertical line should be compared with the best estimates of these parameters (posterior mean) marked by either a solid circle or dashed vertical line. Table 2 summarizes the posterior mean, posterior standard deviation, and lower and upper bounds for the 95% HPD interval for the source parameters.

It is clear from an examination of Figure 4 and of Table 2 that all the source parameters have been recovered with very good accuracy. More specifically, the distance between the actual source location and the best estimate (posterior mean) of the source location is about 44 km (see Figure 5), and the recovered emission rate is within 20% of the actual emission rate. Note that the accuracy in the inferred source location is roughly a factor of ten better than that obtained using the standard measurement model (which does not use multipliers to try to compensate for the unknown model errors). More importantly, the precision of the source parameter estimates (namely, the 95% credible intervals) for
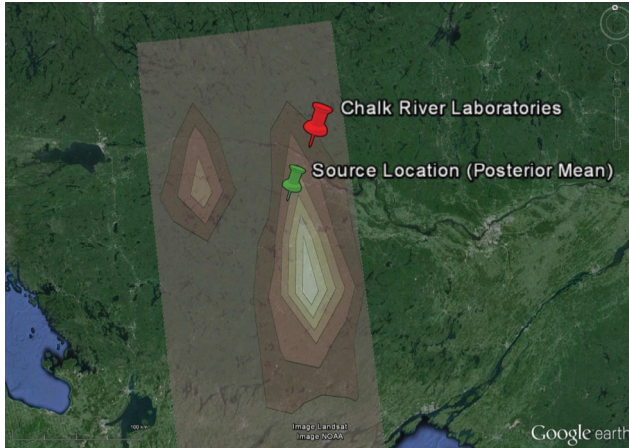
FIGURE 5: Two-dimensional marginal posterior distribution of the source location georeferenced on a Google Earth image. The marginal distribution is for the alternative measurement model which employs multipliers to compensate for the unknown model error structure.

the alternative measurement model which utilizes multipliers to compensate for model errors contains the actual (true) values for the source parameters, in stark contrast to the reconstruction using the standard measurement model (cf. the inferred values for the source parameters in Table 1 with those in Table 2).

In the practical real-world example of source reconstruction considered herein using activity concentration data obtained from the IMS radionuclide network, it is extremely difficult to correctly specify *a priori* both the magnitude (or scale) of the model error and the model error structure. In light of this difficulty, it is apparent from this example that incorporating multipliers (scale factors) with the predicted concentrations to attempt to compensate for the model error can improve significantly the quality of the source reconstruction. As an indication of the complexity in the model error structure for the current problem, Figure 6 exhibits the one-dimensional marginal posterior distributions for the various multipliers m_J ($J = 1, 2, \ldots, 8$). An examination of Figure 6 provides an indication of the complexity in the heteroskedastic model error structure, which makes it difficult to provide *a priori* estimates for the model error for use in the source reconstruction. Note that the distributions for the multipliers associated with some of the predicted concentrations are quite broad, indicating that the model errors for these predicted concentrations are not well determined. In other cases, the distributions for the multipliers (e.g., m_2 and m_3) are quite narrow, implying that the data contain reasonable information to constrain the model errors associated with these multipliers fairly well.

The true values for the multipliers are not known. However, we can obtain an independent estimate for the values of the multipliers in the current example because the actual source parameters are known. We can use the

actual source parameters $\boldsymbol{\theta}_s^*$ to predict the activity concentration $\overline{C}_J(\boldsymbol{\theta}_s^*)$ that would be expected at the various observation stations. With this information, we can provide an independent point estimate for the multipliers as $\widehat{m}_J = d_J/C_J(\boldsymbol{\theta}_s^*)$, $J = 1, 2, \ldots, 8$. These "actual" values for the multipliers are shown as the solid vertical lines in Figure 6, which should be compared with best estimates of the multipliers obtained as the mean of the posterior distributions for m_J. A visual inspection shows that the best estimates of the multipliers are consistent with the "actual" values for the multipliers obtained from the indicated point estimates. Furthermore, the width of the posterior distribution of the various multipliers is seen to enclose the "actual" values for the multipliers. Finally, even though some of the multipliers are highly uncertain, inclusion of multipliers to compensate for the model errors does lead to significantly improved estimates for the source parameters (both in accuracy and in precision).

5. Conclusions

In this study, a Bayesian inferential methodology has been applied to the problem of source reconstruction for real-world activity concentration data measured by an operational network of sensors (more particularly by the IMS radionuclide network that is maintained under the auspices of the United Nations CTBT Organization). This methodology for source reconstruction has been applied to a difficult situation, namely, reconstruction of the source characteristics from the CRL release using only a small number of activity concentration measurements (8 measurements) obtained from only three sampling stations. The problem involved utilization of an operational backward LS model for long-range transport by the atmosphere on a continental scale.

The principal difficulty in the reconstruction lay in providing the correct *a priori* specification of the model error for the various predicted concentrations associated with the measured concentrations used for the source parameter recovery. A naïve specification for the model error gave a source reconstruction that was quite reasonable in terms of the accuracy in the recovery of the location and emission rate, but the precision in the estimates was generally poor in the sense that the reported uncertainty bounds in the recovery of the source parameters did not include the actual (true) values for these parameters. This led to an alternative measurement model which incorporated multipliers (scale factors) with the predicted concentrations in order to compensate for the model errors. The application of this alternative model was shown to yield significantly improved estimates for the source parameters, both in terms of their accuracy and their precision.

Although Bayesian probability theory offers a coherent and rational approach for source reconstruction, its application to real-world problems using real sensor networks and operational dispersion models will require a better understanding of both the scale and structure of the model error in the predicted concentrations. It is anticipated that the proper incorporation of information about the underlying model error in conjunction with Bayesian inference employing
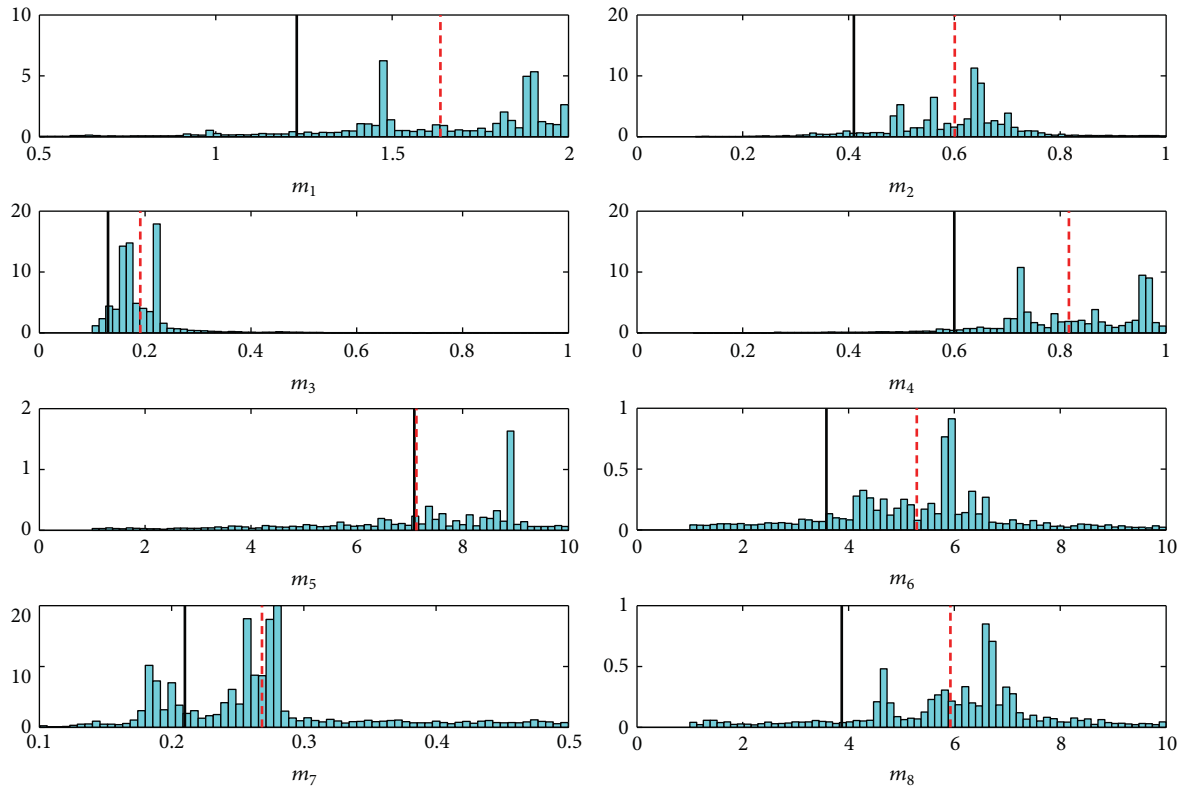
FIGURE 6: One-dimensional marginal posterior distribution of the multipliers m_J ($J = 1, 2, \ldots, 8$) used to compensate for the model error in the problem. The "actual" values of the multipliers are indicated with the solid vertical line which should be compared with the best estimates of the multipliers (posterior means).

state-of-the-science MCMC sampling algorithms (such as the MT-DREAM$_{(ZS)}$ algorithm) will provide the most flexible framework for source reconstruction.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] D. A. Shea and S. A. Lister, "The BioWatch program: detection of bioterrorism (online)," Congressional Research Service Report RL 35152, 2003, http://fas.org/sgp/crs/terror/RL32152.html.

[2] CTBTO Preparatory Commission, *The Global Verification Regime and the International Monitoring System*, Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO), 2001.

[3] L. Robertson and C. Persson, "Attempts to apply four dimensional data assimilation of radiological data using the adjoint technique," *Radiation Protection Dosimetry*, vol. 50, no. 2–4, pp. 333–337, 1993.

[4] L. Robertson and J. Langner, "Source function estimate by means of variational data assimilation applied to the ETEX-I tracer experiment," *Atmospheric Environment*, vol. 32, no. 24, pp. 4219–4225, 1998.

[5] L. C. Thomson, B. Hirst, G. Gibson et al., "An improved algorithm for locating a gas source using inverse methods," *Atmospheric Environment*, vol. 41, no. 6, pp. 1128–1134, 2007.

[6] M. Bocquet, "Reconstruction of an atmospheric tracer source using the principle of maximum entropy—I: theory," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 610, pp. 2191–2208, 2005.

[7] C. T. Allen, G. S. Young, and S. E. Haupt, "Improving pollutant source characterization by better estimating wind direction with a genetic algorithm," *Atmospheric Environment*, vol. 41, no. 11, pp. 2283–2289, 2007.

[8] A. Keats, E. Yee, and F.-S. Lien, "Bayesian inference for source determination with applications to a complex urban environment," *Atmospheric Environment*, vol. 41, no. 3, pp. 465–479, 2007.

[9] F. K. Chow, B. Kosović, and S. Chan, "Source inversion for contaminant plume dispersion in urban environments using building-resolving simulations," *Journal of Applied Meteorology and Climatology*, vol. 47, no. 6, pp. 1533–1572, 2008.

[10] I. Senocak, N. W. Hengartner, M. B. Short, and W. B. Daniel, "Stochastic event reconstruction of atmospheric contaminant dispersion using Bayesian inference," *Atmospheric Environment*, vol. 42, no. 33, pp. 7718–7727, 2008.

[11] E. Yee, F.-S. Lien, A. Keats, and R. D'Amours, "Bayesian inversion of concentration data: source reconstruction in the adjoint representation of atmospheric diffusion," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96, no. 10-11, pp. 1805–1816, 2008.

[12] E. Yee, "Theory for reconstruction of an unknown number of contaminant sources using probabilistic inference," *Boundary-Layer Meteorology*, vol. 127, no. 3, pp. 359–394, 2008.

[13] E. Yee, "Probability theory as logic: data assimilation for multiple source reconstruction," *Pure and Applied Geophysics*, vol. 169, no. 3, pp. 499–517, 2012.

[14] E. Yee, "Source reconstruction: a statistical mechanics perspective," *International Journal of Environment and Pollution*, vol. 48, no. 1–4, pp. 203–213, 2012.

[15] E. Yee, "Inverse dispersion for an unknown number of sources: model selection and uncertainty analysis," *ISRN Applied Mathematics*, vol. 2012, Article ID 465320, 20 pages, 2012.

[16] R. Cox, "Probability, frequency, and reasonable expectation," *American Journal of Physics*, vol. 14, no. 1, pp. 1–13, 1946.

[17] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, 2003.

[18] G. I. Marchuk, "Formulation of the theory of perturbations for complicated models," *Applied Mathematics and Optimization*, vol. 2, no. 1, pp. 1–33, 1975.

[19] F. X. Le Dimet and O. Talagrand, "Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects," *Tellus A*, vol. 38, no. 2, pp. 97–110, 1986.

[20] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, CRC Press (Chapman and Hall), 1991.

[21] A. Gelman, J. B. Carlin, H. S. Stern, and B. R. Rubin, *Bayesian Data Analysis*, CRC Press, 2nd edition, 2004.

[22] V. Savchuk and C. P. Tsokos, *Bayesian Theory and Methods*, Springer, 2014.

[23] K. -V. Yuen, *Bayesian Methods for Structural Dynamics and Civil Engineering*, John Wiley and Sons, 2010.

[24] E. Laloy and J. A. Vrugt, "High-dimensional posterior exploration of hydrologic models using multiple-try DREAM(ZS) and high-performance computing," *Water Resources Research*, vol. 48, no. 1, Article ID W01526, 2012.

[25] J. S. Liu, F. Liang, and W. H. Wong, "The multiple-try method and local optimization in Metropolis sampling," *Journal of the American Statistical Association*, vol. 95, no. 449, pp. 121–134, 2000.

[26] A. G. Gelman and D. B. Rubin, "Inference from iterative simulation using multiple sequences," *Statistical Science*, vol. 7, no. 4, pp. 457–472, 1992.